



WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft  
*The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics*

Bönte, G.

Working Paper

# Evaluierung adaptiver Hochrechnungsverfahren

Diskussionspapiere // Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für  
Statistik und Ökonometrie, No. 05/1995

**Provided in cooperation with:**

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Suggested citation: Bönte, G. (1995) : Evaluierung adaptiver Hochrechnungsverfahren,  
Diskussionspapiere // Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Statistik  
und Ökonometrie, No. 05/1995, <http://hdl.handle.net/10419/29604>

**Nutzungsbedingungen:**

Die ZBW räumt Ihnen als Nutzerin/Nutzer das unentgeltliche,  
räumlich unbeschränkte und zeitlich auf die Dauer des Schutzrechts  
beschränkte einfache Recht ein, das ausgewählte Werk im Rahmen  
der unter

→ <http://www.econstor.eu/dspace/Nutzungsbedingungen>  
nachzulesenden vollständigen Nutzungsbedingungen zu  
vervielfältigen, mit denen die Nutzerin/der Nutzer sich durch die  
erste Nutzung einverstanden erklärt.

**Terms of use:**

*The ZBW grants you, the user, the non-exclusive right to use  
the selected work free of charge, territorially unrestricted and  
within the time limit of the term of the property rights according  
to the terms specified at*

→ <http://www.econstor.eu/dspace/Nutzungsbedingungen>  
*By the first use of the selected work the user agrees and  
declares to comply with these terms of use.*



Leibniz-Informationszentrum Wirtschaft  
Leibniz Information Centre for Economics

Mitglied der



# Inhaltsverzeichnis

<b>1. Einführung und Motivation</b>	<b>2</b>
<b>2. Die robusten Schätzverfahren</b>	<b>3</b>
2.1. M-Regressionsschätzung . . . . .	3
2.2. Anwendung der Schätzverfahren . . . . .	4
<b>3. Adaptierende Schätzverfahren</b>	<b>5</b>
3.1. Adaptierende Regressionsverfahren nach Yuh und Hogg . . . . .	6
3.2. Das adaptierende Regressionsverfahren nach Moberg, Ramberg und Randles . . . . .	9
3.3. Das adaptierende Regressionsverfahren nach Kappenman und das teilweise adaptierende Regressionsverfahren nach McDonald und Newey . . . . .	11
<b>4. Das Bootstrap-Verfahren</b>	<b>13</b>
<b>5. Das Datenmaterial und die Hochrechnung</b>	<b>14</b>
<b>6. Analyseergebnisse</b>	<b>15</b>
<b>7. Wahl des Startwertes</b>	<b>17</b>
<b>8. Zusammenfassung</b>	<b>19</b>

# 1. Einführung und Motivation

Die Konjunkturforschungsstelle an der Eidgenössischen Technischen Hochschule zu Zürich (KOF/ETH) führt seit 1967 eine jährliche Investitionsbefragung bei den schweizerischen Unternehmen durch. Ziel dieser Befragung ist es, nach Branchen gegliedert Informationen über die effektiv getätigten und für die im Berichtsjahr geplanten Bau- und Ausrüstungsinvestitionen zu erhalten.

Die Befragung erfolgt in schriftlicher Form. Durch die Hochrechnung der Stichprobenergebnisse wird versucht,

- (i) anhand der Stichprobe auf das Investitionsverhalten der nicht an der Befragung teilnehmenden restlichen Firmen zu schließen, um eine Vorstellung über das Gesamtniveau der privaten Investitionen in der Schweiz zu erhalten und
- (ii) die amtliche Statistik durch Hochrechnung der Investitionsangaben auf Branchenebene zu ergänzen.

Für die Hochrechnung wird eine Größe benötigt, die für alle Unternehmen verfügbar ist und in einem möglichst engen Zusammenhang zu den Investitionen steht. Hierfür ist auf die Beschäftigtenzahlen zurückgegriffen worden, die aus einer Betriebsstättenzählung des Jahres 1985 entnommen worden sind.

Es gelte das einfache Regressionsmodell

$$Y_{n \times 1} = X_{n \times 1} \beta_{1 \times 1} + U_{n \times 1}, \quad (1)$$

mit

$Y$ : Vektor der getätigten Investitionen,

$X$ : Vektor der Beschäftigtenzahlen,

$\beta$ : Zu schätzender Parameter,

$U$ : Vektor der Störgrößen der Regression.

Unter Berücksichtigung einer nichtnegativen Funktion der Beschäftigtenzahlen  $v(X_i)$ , die die Spezifikation verschiedener Annahmen über den Fehler der Regression erlaubt, lautet das Modell für das  $i$ -te Unternehmen

$$Y_i = X_i \beta + \sqrt{v(X_i)} U_i. \quad (2)$$

Zur Ermittlung von  $\beta$  stellt der traditionelle Regressionsschätzer (Kleinst-Quadrat-Schätzer) ein übliches Verfahren dar. Seine Problematik liegt jedoch darin, daß einige ungewöhnliche Residuen (Ausreißer) einen bei weitem größeren Einfluß auf den Schätzwert ausüben können als andere Beobachtungen.

Die Erforschung und Entwicklung "robuster" Verfahren ist im wesentlichen auf diese Problematik des Kleinst-Quadrat-Schätzers zurückzuführen. Die robusten Methoden sollen zwei Ziele sicherstellen:

- (i) Schutz gegen ungewöhnliche Datenbefunde (Ausreißer).
- (ii) Gute Effizienz, wenn die Daten aus dem idealen Normalverteilungsmodell (Gaußmodell) oder aus einer Menge alternativer Modelle stammen.

Eine Weiterentwicklung der robusten stellen die “adaptierenden“ Verfahren dar, deren Ziel in der Bestimmung der Verteilung der Daten liegt, um daraus folgend eine Schätzung abzuleiten.

Diese genannten Verfahren, Kleinst-Quadrat, robuste und adaptierende, werden im folgenden auf die Daten des Investitionstestes zur Hochrechnung der Investitionen angewendet.

In den Abschnitten 2 und 3 werden die verwendeten robusten und adaptierenden Schätzverfahren vorgestellt. In Abschnitt 4 wird auf das Bootstrapverfahren eingegangen, das zur Untersuchung der statistischen Eigenschaften der Schätzverfahren herangezogen wird. Die Abschnitte 5 bis 7 dienen der Vorstellung der Analyseergebnisse.

## 2. Die robusten Schätzverfahren

### 2.1. M-Regressionsschätzung

Ein M-Schätzer für  $\beta$  des Modells (1), basierend auf einer Zielfunktion  $\rho(t)$  und den Daten  $(X_1, Y_1), \dots, (X_n, Y_n)$  ist der Wert von  $\beta$ , bezeichnet durch  $\hat{\beta}_M$ , der den Ausdruck

$$\sum_{i=1}^n \rho(Y_i - X_i \beta) \quad (3)$$

minimiert.

Sofern  $\rho(Y_i - X_i \beta)$  nach  $\beta$  differenzierbar ist, kann  $\hat{\beta}_M$  auch als Lösung von

$$\sum_{i=1}^n \psi(Y_i - X_i \beta) X_i = 0 \quad (4)$$

berechnet werden, mit  $\psi$  als Ableitung von  $\rho$ .  $\psi$  wird Scores- oder Gewichtsfunktion genannt. Die Wahl von  $\psi$  bestimmt entscheidend die Eigenschaften eines Schätzers, wie im folgenden zu sehen sein wird.

Die wohl am meisten in der Literatur untersuchten robusten Regressionsverfahren sind die nach Huber und Hampel.

Die Huber-Verteilung folgt in der Mitte einer Normal- und an den Rändern einer Doppelt-Exponentialverteilung, so daß die Scorefunktion lautet:

$$\psi(u) = -kI_{(-\infty, -k)}(u) + uI_{[-k, k]}(u) + kI_{(k, \infty)}(u). \quad (5)$$

Jenseits der Tuningkonstanten  $k$  bleibt die Scoresfunktion konstant. Die Wahl von  $k$  beeinflußt ebenso die Effizienz des Schätzers, wenn die Daten normalverteilt sind, wie auch den Schutz gegen ungewöhnliche Beobachtungen. Mit sinkendem  $k$  steigt die Robustheit des Schätzers bei gleichzeitig abnehmender Effizienz und umgekehrt.

Es ist im Rahmen dieser Untersuchung  $k = 2$  gewählt worden, was einer Effizienz von 99.1% entspricht, wenn die Daten aus einer Normalverteilung stammen.

Bei dem Schätzer von Hampel besteht die Gewichtsfunktion aus drei Teilen:

$$\psi(u) = \begin{cases} u & \text{für } |u| \leq a \\ a \cdot \text{sign}(u) & \text{für } a < |u| \leq b \\ a \frac{c \cdot \text{sign}(u) - u}{c - b} & \text{für } b < |u| \leq c \\ 0 & \text{für } |u| > c \end{cases} \quad (6)$$

Die Parameter  $a$ ,  $b$  und  $c$  stellen Tuningkonstanten dar, deren Wahl sich wie zuvor beim Huber-Schätzer direkt auf die Effizienz auswirkt, wenn die Daten aus dem Gauß-Modell stammen, wie auch auf das Maß an Sicherheit gegen ungewöhnliche Beobachtungen.

In dieser Untersuchung ist  $a = 2.5$ ,  $b = 4.5$  und  $c = 9.5$  gewählt worden, was einer Effizienz von 95.6% entspricht, wenn die Daten einer Normalverteilung folgen.

Die  $\psi$ -Funktion wird als "redescending" bezeichnet, da die Gewichte ab dem dritten Abschnitt bis auf Null zurückgehen.

## 2.2. Anwendung der Schätzverfahren

Die Regressionsschätzer nach Huber und Hampel und die später vorgestellten adaptierenden Schätzverfahren lassen sich iterativ bestimmen. Es wird zwischen "voll iterierten" und "Ein-Schritt-Schätzern" unterschieden.

Bei voll iterierten Schätzverfahren läuft der Algorithmus bis zur Konvergenz. Ein solcher Algorithmus wurde von Huber (1977) vorgeschlagen:

$$\hat{\beta}^{m+1} = \hat{\beta}^m + \hat{\sigma}(X'X)^{-1}X'\psi\left(\frac{Y - X\hat{\beta}^m}{\hat{\sigma}}\right). \quad (7)$$

$\hat{\sigma}$  stellt einen Skalenschätzer dar. Um Skaleninvarianz zu erhalten, werden sämtliche Residuen durch diesen dividiert.

Ein Beispiel für einen solchen Skalenschätzer ist der *MAD* (Median Absolute Deviations), der sich darstellt als

$$\hat{\sigma} = \frac{1}{0.6745} \text{med}\{|Y_i - X_i\hat{\beta}^0 - \text{med}\{Y_i - X_i\hat{\beta}^0\}|\}, \quad (8)$$

wobei  $\hat{\beta}^0$  ein Startschätzer und 0.6745 der durchschnittliche Wert des MAD für Stichproben aus dem Gaußmodell sind.

In der Literatur wird überwiegend vorgeschlagen, als Startschätzer  $\hat{\beta}^0$  den *Least-Absolute-Residuals*-Schätzer (LAR) zu verwenden, der die Lösung des Ausdrucks

$$\hat{\beta}_{LAR} = \min_{\beta} \sum_{i=1}^n \left| \frac{Y_i - X_i \beta}{\hat{\sigma}} \right| \quad (9)$$

ist.

Die Gewichtsfunktion des LAR-Schätzer hat die Form

$$\psi(u) = \text{sign}(u). \quad (10)$$

Der LAR-Schätzer scheint sehr attraktiv für robuste Regressionsverfahren zu sein, denn er weist großen Residuen ebensoviel Gewicht zu wie kleinen.

In den folgenden Anwendungen kommt dennoch neben dem LAR- auch der Kleinst-Quadrat-Schätzer als Startwert zur Anwendung. Huber (1973) bemerkt, daß dieser trotz der negativen Eigenschaften bezüglich Ausreißern als Anfangsschätzer durchaus geeignet sein könne.

Ein Ziel dieser Untersuchung ist es, demnach auch festzustellen, welcher Startwert bessere Eigenschaften besitzt.

Im Gegensatz zu den voll iterierten werden Ein-Schritt-Schätzer bestimmt, indem von einem Startwert ausgehend der Algorithmus nur einmal durchlaufen wird:

$$\hat{\beta}_{one-step} = \hat{\beta}_0 - \frac{\sum_{i=1}^n \psi(u_i)}{\sum_{i=1}^n \psi'(u_i)}. \quad (11)$$

Obwohl von den Eigenschaften her asymptotisch äquivalent für  $n \rightarrow \infty$  zu den voll iterierten Schätzern, ist es offensichtlich, daß diese sehr sensibel gegenüber der Wahl des Anfangsschätzers sind.

Ein weiteres Ziel dieser Untersuchung ist es daher, Empfehlungen zu geben, ob sich der größere Rechenaufwand für die Bestimmung der voll iterierten Schätzverfahren hinsichtlich eines besseren Ergebnisses rentiert, oder ob auch die Ein-Schritt-Schätzer konkurrenzfähige Resultate liefern können.

### 3. Adaptierende Schätzverfahren

Die (grob) adaptierenden Verfahren versuchen, die Residuen einer Initialregression in Verteilungsklassen hinsichtlich Schwere der Tails und Schiefe einzuteilen, wobei für jede Klasse eine individuelle Gewichtsfunktion angeboten wird.

In der Literatur sind vor allem bezüglich der Lokalisationsschätzung eine Fülle von Vorschlägen gemacht worden (vgl. z.B. Hogg und Randles (1972) oder Hogg und Russel (1983)).

Für die Anwendung bei der Regressionsschätzung soll hier auf die Arbeiten von Yuh und Hogg (1988), von Moberg, Ramberg und Randles (1980), von Kappenman (1988) und von McDonald und Newey (1988) eingegangen werden.

#### 3.1. Adaptierende Regressionsverfahren nach Yuh und Hogg

Yuh und Hogg (1988) gehen bei ihren Vorschlägen zur adaptierenden Regression so vor, daß zunächst die Residuen  $u_1, \dots, u_n$  der Anfangsregression der Größe nach geordnet werden (bezeichne  $z_1, \dots, z_n$  diese geordneten Residuen). Diese geordneten Residuen werden herangezogen, um Aussagen über Schiefe und Stärke der Tails ihrer Verteilung zu treffen.

Um auf Symmetrie einer Verteilung zu testen, wird die Statistik

$$H_1 = \frac{U_{0.05} - M_{0.40}}{M_{0.40} - L_{0.05}} \quad (12)$$

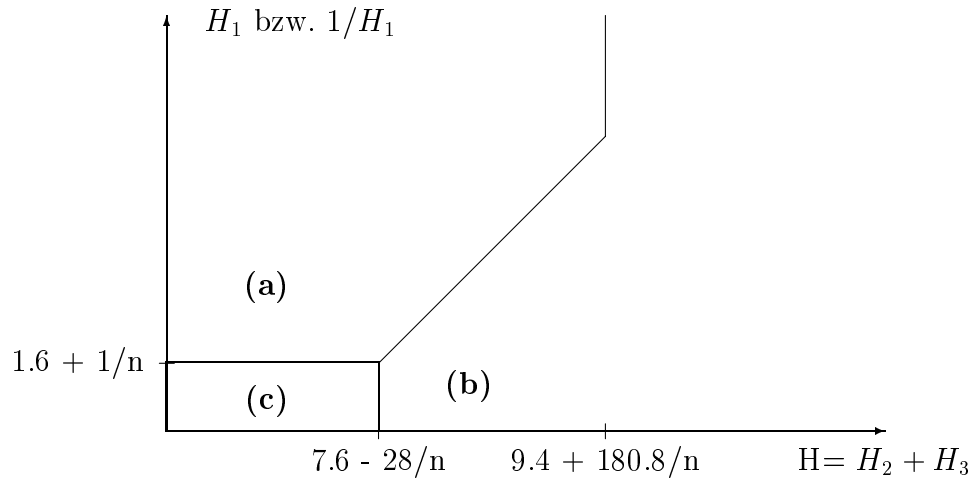
vorgestellt.

Für einen Test auf Stärke der Tails wird die Statistik  $H = H_2 + H_3$  eingeführt.  $H_2$  und  $H_3$  sind definiert als

$$\begin{aligned} H_2 &= \frac{E_{0.10} - B_{0.10}}{D_{0.38} - C_{0.38}} \\ H_3 &= \frac{U_{0.02} - L_{0.02}}{D_{0.48} - C_{0.48}}, \end{aligned} \quad (13)$$

wobei  $L, B, C, D, E, U$  die Stichprobe in sechs Teile partitionieren. Zum Beispiel bezeichne  $L_{0.05}$  den Durchschnitt der ersten  $[n \cdot 0.05]$  Ordnungsstatistiken und  $U_{0.05}$  den Durchschnitt der größten  $[n \cdot 0.05]$  Ordnungsstatistiken.  $M_{0.40}$  stellt das  $1 - 2\delta$ -getrimmte Mittel mit  $\delta = 0.30$  dar.

Das Klassifikationsschema hinsichtlich Tailsstärke und Schiefe hat dann folgende Gestalt:



(a): starke rechtsschiefe (bzw. linksschiefe)

Verteilung mit schweren Tails

(b): symmetrische Verteilung mit schweren Tails.

(c): symmetrische Verteilung mit leichten Tails.

Man kann durch dieses Klassifikationsschema jedoch Verteilungen nur bezüglich Schwere der Tails und Rechtsschiefe untersuchen.

Liegen linksschiefe Verteilungen vor, so muß die Statistik  $1/H_1$  anstatt  $H_1$  verwendet werden.

Es kann also nicht gleichzeitig die Schwere der Tails und die Schiefe in beide Richtungen untersucht werden. Um einen Automatismus der Klassifikation via Computer zu erreichen, muß dieser Algorithmus noch um die Information der Art der Schiefe erweitert werden.

Hierfür bietet sich das Schiefemaß an, das auf die geordneten Residuen angewendet wird

$$\omega = \frac{\text{Mittelwert} - \text{Median}}{\text{Standardabweichung}}. \quad (14)$$

Alternativ könnte auch der Schiefekoeffizient  $\mu_3/\sigma^3$  verwendet werden, jedoch hat  $\omega$  den Vorteil, daß das dritte zentrale Moment nicht berechnet werden muß.

Die Erweiterung der Klassifikation sieht demnach wie folgt aus:

Wird die Verteilung als "schief" erkannt, so bilde  $\omega$ .

Wenn  $\omega > 0$ : Verteilung wird als linksschief eingestuft.

Wenn  $\omega < 0$ : Verteilung wird als rechtsschief eingestuft.

Bilde auf Basis dieses Ergebnisses die im folgenden vorgestellten Gewichtsfunktionen.



Es werden dabei drei verschiedene Varianten eines adaptierenden Schätzers entwickelt.

Der erste ist der adaptierende Schätzer basierend auf der Student t-Verteilung (adaptive t-weight estimator).

Die Gewichtsfunktion dieses Schätzers lautet

$$\psi(z) = \frac{z}{1 + \frac{z^2}{r}}, \quad (15)$$

mit  $r$  als Anzahl der Freiheitsgrade.

Als Repräsentant einer Verteilung mit leichten bzw. schweren Tails wird die t-Verteilung mit elf Freiheitsgraden bzw. drei Freiheitsgraden, für rechtsschiefe Verteilungen eine Gewichtsfunktion der Form

$$\psi(z) = \frac{z}{1 + 5 \cdot (z + 0.5)^2} \quad (16)$$

herangezogen.

Für "linksschiefe" Verteilungen wird der Ausdruck  $(z + 0.5)$  durch  $(z - 0.5)$  ersetzt.

Als zweites Verfahren wird der adaptierende Schätzer mit linearen Splines (adaptive estimator using linear splines) vorgestellt.

Wenn die Klassifikation eine Verteilung mit leichten Tails ergibt, wird

$$\psi(z) = z \quad (17)$$

verwendet, d.h. der Schätzer ist identisch mit dem Kleinst-Quadrat-Schätzer. Für den Fall einer Verteilung mit schweren Tails, wird

$$\psi(z) = \begin{cases} \frac{z}{1.5} & \text{für } 0 < z < 1.5 \\ \max(0, 1 - 0.05(z - 1.5)) & \text{für } z \geq 1.5 \end{cases} \quad (18)$$

$$\text{mit } \psi(-z) = -\psi(z)$$

vorgeschlagen.

Für den Fall einer schiefen Verteilung wird folgende Gewichtsfunktion angenommen:

$$\psi(z) = \begin{cases} \min(0, -1.5z - 2.5) & \text{für } z \leq -1 \\ z & \text{für } -1 < z < 0.01 \\ 0.01 & \text{für } z \geq 0.01 \end{cases} \quad (19)$$

Weiterhin schlagen Yuh und Hogg einen *0-Schritt-Schätzer* als Anfangsschätzer für einen Ein-Schritt-Schätzer vor, nachdem die Klassifikation durch die vorhergehende Schätzung bestimmt wurde. Zu begründen ist dieses Vorgehen durch die Sensitivität der Ein-Schritt-Schätzer in Bezug auf die Initialschätzung. Es soll also gewährleistet werden, daß ein der jeweiligen Situation angemessener Startschätzer zur Anwendung kommt.

Der 0-Schritt-Schätzer  $\beta_{0-step}$  hat folgende Gestalt:

$$\hat{\beta}_{0-step} = \begin{cases} \hat{\beta}_{Kleinst-Quadrat} & \text{leichte Tails} \\ \hat{\beta}_{LAR} & \text{schwere Tails} \\ \hat{\beta}_{L_{0.25}} & \text{rechtsschiefe Verteilung} \\ \hat{\beta}_{L_{0.75}} & \text{linksschiefe Verteilung,} \end{cases} \quad (20)$$

wobei  $\hat{\beta}_{L_p}$  der Schätzer ist, der folgenden Ausdruck minimiert:

$$\sum_{R_1} p \cdot |Y_i - X_i\beta| + \sum_{R_2} (1-p) \cdot |Y_i - X_i\beta|, \quad (21)$$

mit  $R_1 = \{i : Y_i \geq X_i\beta\}$ ,  $R_2 = \{i : Y_i < X_i\beta\}$  und  $0 < p < 1$ .

Es ist direkt ersichtlich, daß der LAR-Schätzer ein Spezialfall von  $\hat{\beta}_p$  mit  $p = 0.5$  ist.

### 3.2. Das adaptierende Regressionsverfahren nach Moberg, Ramberg und Randles

Moberg, Ramberg und Randles (1980) (im folgenden als Schätzer nach Moberg bezeichnet) verwenden mit fünf möglichen Klassen der Residuenverteilung eine feinere Einteilung als Yuh und Hogg.

Es werden die Statistiken  $Q_3$  zur Untersuchung der geordneten Residuen einer Anfangsschätzung hinsichtlich Schiefe und  $Q_4$  bezüglich Tailsstärke eingeführt. Sie sind definiert als

$$\begin{aligned} Q_3 &= \frac{\overline{U}(0.05) - \overline{M}(0.50)}{\overline{M}(0.50) - \overline{L}(0.05)} \\ Q_4 &= \frac{\overline{U}(0.05) - \overline{L}(0.05)}{\overline{U}(0.50) - \overline{L}(0.50)}, \end{aligned} \quad (22)$$

mit

$\overline{U}(\gamma)$ : Durchschnitt der größten  $[n \cdot \gamma]$  Ordnungsstatistiken.

$\overline{M}(\gamma)$ : Durchschnitt der mittleren  $[n \cdot \gamma]$  Ordnungsstatistiken.

$\bar{L}(\gamma)$ : Durchschnitt der kleinsten  $[n \cdot \gamma]$  Ordnungsstatistiken.

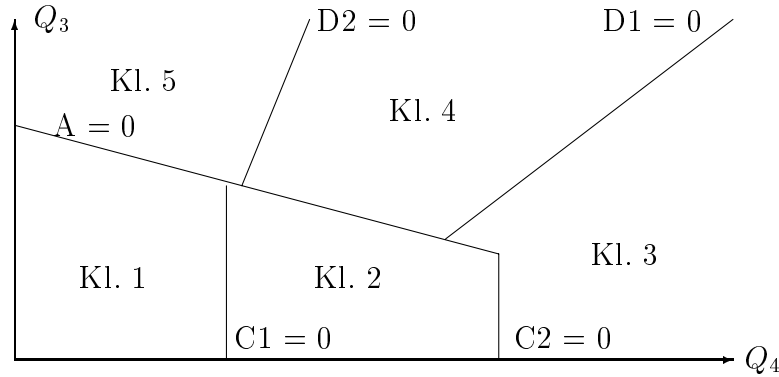
Auf Basis von  $Q_3$  und  $Q_4$  werden die Klassifikationsstatistiken  $A$ ,  $C1$ ,  $C2$ ,  $D1$ , und  $D$  gebildet:

$$\begin{aligned} A &= Q_4 - (1.7 - 9/n)Q_3 - 5.46 + 21.6/n \\ C1 &= Q_4 - 2.011 + 2/n \\ C2 &= Q_4 - 3.011 + 11/n \\ D1 &= Q_4 - (0.75 - 1/n)Q_3 - (1.65 - 2/n) \\ D2 &= Q_4 - (0.5 - 1/n)Q_3 - (1.25 - 1/n) \end{aligned}$$

Die Stichprobe wird nun den Klassen zu folgenden Bedingungen zugewiesen:

Falls  $A \leq 0$ ,  $C1 < 0$ , wähle Klasse 1.  
 Falls  $A \leq 0$ ,  $C1 \geq 0$ ,  $C2 < 0$ , wähle Klasse 2.  
 Falls  $A \leq 0$ ,  $C1 \geq 0$ ,  $C2 \geq 0$ , wähle Klasse 3.  
 Falls  $A > 0$ ,  $D1 > 0$ , wähle Klasse 3.  
 Falls  $A > 0$ ,  $D1 \leq 0$ ,  $D2 > 0$ , wähle Klasse 4.  
 Falls  $A > 0$ ,  $D1 \leq 0$ ,  $D2 \leq 0$ , wähle Klasse 5.

Das Klassifikationsschema hat folgende Gestalt:



Die Gewichtsfunktionen der einzelnen Klassen lauten:

$$\psi(z) = \begin{cases} z^3 & \text{für die Klasse 1.} \\ z & \text{für die Klasse 2 (Kleinst-Quadrat-Schätzer).} \\ \frac{2.25z}{z^2 + 0.51s} & \text{für die Klasse 3.} \\ \frac{5.10z}{(z + 0.74s)^2 + 0.87s^2} & \text{für die Klasse 4.} \\ \frac{3.30z}{(z + 0.54s)^2 + 0.23s^2} & \text{für die Klasse 5.} \end{cases}$$

mit  $s$  als Skalenschätzer.

Ebenso wie bei Yuh und Hogg kann die Schiefe nur in eine Richtung hin untersucht werden. Daher erfolgt auch hier die Erweiterung des Vorgehens mit der vorgestellten Statistik  $\omega$ . Wenn eine linksschiefe Verteilung ermittelt wird, so wird anstelle von  $Q_3$  mit  $Q_3^* = 1/Q_3$  gearbeitet. Es werden dann zusätzlich die Klassen 6 und 7 gebildet mit  $\psi(-z)$  der Gewichtsfunktionen der Klassen 4 bzw. 5.

Im Gegensatz zum Verfahren nach Yuh werden hier auch Verteilungen mit leichteren Tails als die der Normalverteilung betrachtet. So bemerkt auch Anscombe (1967), daß zu wenig Aufmerksamkeit den Verteilungen mit leichten Tails gewidmet werde, obwohl dieser Fall durchaus häufig in der Praxis anzutreffen sei.

### 3.3. Das adaptierende Regressionsverfahren nach Kappenman und das teilweise adaptierende Regressionsverfahren nach McDonald und Newey

Ebenso wie Moberg geht auch Kappenman (1988) auf Verteilungen ein, die leichtere Tails als die der Normalverteilung besitzen.

Er bestimmt die Verteilung der Störterme, indem er als mögliche Verteilungsmodelle die  $t$ -, die Doppelt-Exponential- (Laplace-) <sup>1</sup> und die Gamma-Verteilung unterstellt. Kappenman greift dann auf die Maximum-Likelihood-Schätzer der bedingten Verteilung  $f(y|x)$  zurück, um diese als Klassifikationsstatistiken zu verwenden.

In einem detailliert angegebenen Algorithmus werden nun die verschiedenen Likelihood-Funktionen miteinander verglichen, um somit das Verteilungsmodell der Residuen zu bestimmen und daraus den Schätzer zu ermitteln.<sup>2</sup>

McDonald und Newey schlagen das "teilweise adaptierende" Schätzverfahren vor. Es ist dem nach Kappenman sehr ähnlich, denn auch hier erfolgt die Maximierung der Likelihood-Funktionen, wenn auch bei Kappenman die der bedingten Verteilungen.

$\hat{\beta}$  ist die Lösung des Ausdrucks

$$\min_{\beta} \sum_{i=1}^n \rho(Y_i - X_i\beta, \hat{\eta}), \quad (23)$$

---

<sup>1</sup>  $f(x; a, b) = \frac{1}{2\beta} \exp\left(-\frac{|x-a|}{\beta}\right)$ .

<sup>2</sup> Vgl. hierzu ausführlich Bönke (1994).

wobei  $\hat{\eta}$  den Schätzer der Fläche der Verteilung der Residuen bezeichnet, ähnlich z.B. dem Schätzer der Freiheitsgrade der  $t$ -Verteilung.

Als Verteilung der Residuen führen McDonald und Newey die “generalized  $t$  statistic“ (GT) ein, die definiert ist als:

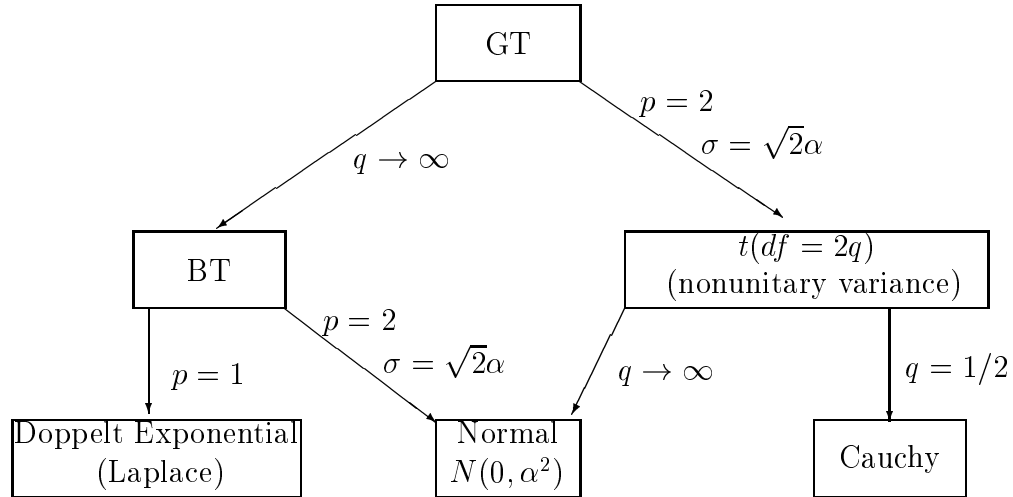
$$f(u; \sigma, p, q) = \frac{p}{2\sigma q^{\frac{1}{p}} B\left(\frac{1}{p}, q\right) \left(1 + \frac{|u|^p}{q\sigma^p}\right)^{p+\frac{1}{p}}} I_{(-\infty, \infty)}(u),$$

mit

$$B\left(\frac{1}{p}, q\right) = \frac{\Gamma\left(\frac{1}{p}\right) \Gamma(q)}{\Gamma\left(\frac{1}{p} + q\right)}. \quad (24)$$

Diese Dichte beherbergt viele bekannte Modelle, denn sowohl Verteilungen mit leichten als auch mit schweren Tails sind durch die Box–Tiao (BT)<sup>3</sup> und die GT-Verteilung erhältlich.

Folgende Übersicht mag diesen Zusammenhang verdeutlichen:



df: Degrees of freedom (Anzahl der Freiheitsgrade) und

BT: Box–Tiao–Verteilung.

Neben dem Regressionsparameter  $\beta$  ist der Vektor  $\eta$  durch  $\hat{\eta}$  zu schätzen. Es können die Residuen der Anfangsschätzung  $\bar{u}_1, \dots, \bar{u}_n$  verwendet werden,

---


$$^3 f(u; \sigma, p) = \frac{p e^{-(|u|/\sigma)^p}}{2\sigma \Gamma\left(\frac{1}{p}\right)}.$$

um mit diesen die Likelihoodfunktion der GT-Verteilung zu maximieren. Der Vektor  $\hat{\eta}$  stellt sich also als Lösung des Ausdruckes

$$\max_{\eta \in \mathcal{H}} \frac{1}{n} \sum \ln f(\bar{u}; \eta)$$

dar, wobei  $\mathcal{H}$  die Menge aller möglichen betrachteten Verteilungen repräsentiert.

## 4. Das Bootstrap-Verfahren

Zur Gütebeurteilung jedes Schätzers erfolgt eine Analyse der Varianz und des Mean-Square-Errors. Diese beiden Größen können durch das Bootstrap-Verfahren ermittelt werden.

Für das einfache lineare Modell (1) kann  $\hat{\beta}$  allgemein als die Lösung des Problems

$$\min_{\beta} D(Y, X\beta) \quad (25)$$

dargestellt werden, mit  $D(\cdot)$  einer Zielfunktion.

Zum Beispiel lautet die Wahl von  $D(Y, X\beta)$  aus (25) bei der Kleinst-Quadrat-Methode

$$D(Y, X\beta) = \sum_{i=1}^n (Y_i - X_i\beta)^2. \quad (26)$$

So wird für jedes vorgestellte Verfahren die jeweilige Funktion von  $D(Y, X\beta)$  eingesetzt.

Aus den resultierenden Residuen  $u_1, u_2, \dots, u_n$  wird eine Bootstrap-Stichprobe der Form

$$Y_i^* = X_i\hat{\beta} + u_i^* \quad i = 1, 2, \dots, n, \quad (27)$$

gebildet, wobei  $u_1^*, u_2^*, \dots, u_n^*$  eine Stichprobe aus den Residuen  $u_1, u_2, \dots, u_n$  bezeichnet. Die Stichprobenelemente seien identisch und unabhängig verteilt.

Für jede Bootstrap-Stichprobe  $Y^*(b)$  wird der Bootstrap-Schätzer  $\hat{\beta}^*(b)$  gebildet

$$\hat{\beta}^*(b) : \min_{\beta} D(Y^*, X\beta).$$

Eine Schätzung des Mean-Square-Errors eines Schätzers erhält man durch Bildung folgenden Ausdrucks:

$$\widehat{MSE}(\hat{\beta}) = \hat{\sigma}_{\hat{\beta}}^2 + \widehat{Bias}(\hat{\beta})^2,$$

wobei gilt:

$$\hat{\sigma}_{\hat{\beta}}^2 = \frac{\sum_{b=1}^B \left( \hat{\beta}^*(b) - \hat{\beta}^*(\cdot) \right)^2}{B-1}$$

und

$$\widehat{Bias}(\hat{\beta}) = \hat{\beta} - \hat{\beta}^*(\cdot),$$

mit

$$\hat{\beta}^*(\cdot) = \sum_{b=1}^B \frac{\hat{\beta}^*(b)}{B}.$$

$B$  bezeichne hierbei die Anzahl der Bootstrap-Stichproben, wobei einer Empfehlung von Efron und Tibshirani (1986) folgend  $B = 200$  gewählt wurde.

## 5. Das Datenmaterial und die Hochrechnung

Das für diese Arbeit zur Verfügung gestellte Datenmaterial umfaßt Daten aus dem Investitionstest der KOF/ETH der Jahre 1986 und 1987 für 25 Branchen. Bei der Hochrechnung ist zu unterscheiden, ob die Unternehmen in Größenklassen (gemessen in Beschäftigtenzahlen) geschichtet vorliegen oder nicht. Darüberhinaus erfolgt eine Schätzung mittels aggregierter Daten, d.h. das gesamte Datenmaterial der Branchen wird zu einer Datei zusammengefaßt, und es erfolgt mit diesen aggregierten Daten eine Schätzung mit und ohne Berücksichtigung der Schichten.

Die Gesamtinvestitionen  $\sum_{i=1}^n Y_i$  lassen sich aufteilen in

$$\sum_{i=1}^n Y_i = \sum_s Y_i + \sum_{\bar{s}} Y_i, \quad (28)$$

wobei der erste Summand die Investitionen der Unternehmungen darstellen, die in der Befragung erfaßt worden sind. Für die Hochrechnung  $T$  gilt es nun, den zweiten Summanden (also die nicht im Test erfaßten Unternehmen) zu prognostizieren. Dies erfolgt unter Hinzunahme des errechneten Regressionsschätzers:

$$T = \sum_s Y_i + \hat{\beta} \sum_{\bar{s}} X_i. \quad (29)$$

Als weiteres Kriterium neben dem Prognosefehler (als die Wurzel aus dem Mean-Square-Error) zur Beurteilung der Güte der jeweiligen Schätzverfahren dient der Vergleich der Hochrechnungsergebnisse mit den Angaben der “Nationalen Buchhaltung“, der Schweizer Volkswirtschaftlichen Gesamtrechnung.<sup>4</sup>

---

<sup>4</sup>Vgl. hierzu ausführlich Bönnte (1994, S. 61–67), und Klein (1989, S. 7–8).

## 6. Analyseergebnisse

In diesem Abschnitt werden die Hochrechnungsergebnisse und der Prognosefehler als Gütekriterien der jeweiligen Schätzverfahren vorgestellt. Die Sollwerte für die Hochrechnung wurden von Klein (1989, S. 16–18) ermittelt und betragen für die Bauinvestitionen des Jahres 1986 9689 Mill. SFr. Es soll hier nur auf die desaggregierten Bauinvestitionen des Jahres 1986 zurückgegriffen werden.<sup>5</sup>

Tabelle 1: Hochrechnungsergebnisse der desaggregierten Bauinvestitionen in Mill. SFr. für das Jahr 1986.

Schätzverfahren	Mit Schichtenbildung		Ohne Schichtenbildung	
	Hochrechnung	Differenz in %	Hochrechnung	Differenz in %
1. KQ	10610.15	9.51	10623.40	9.64
2. LAR	5373.37	-44.54	5894.64	-39.16
Startwert KQ-Schätzer:				
3. Huber	6739.00	-30.45	7118.05	-26.53
4. Huber <sub>one-step</sub>	12346.85	27.43	11043.63	13.98
5. Hampel	6406.60	-33.88	6372.71	-34.23
6. Hampel <sup>1</sup>	8143.52	-15.95	8369.34	-13.62
7. Hampel <sub>one-step</sub>	12634.55	30.40	12064.27	24.52
8. McDonald / Newey	9298.11	-4.03	9780.85	0.95
9. Kappenman	9057.94	-6.51	9202.36	-5.02
10. Yuh-AT	5033.63	-48.05	5165.13	-46.69
11. AT <sub>one-step</sub>	6574.60	-32.14	6900.62	-28.78
12. Yuh-ALS	5070.90	-47.66	5276.27	-45.54
13. ALS <sub>one-step</sub>	9124.73	-5.82	7757.11	-19.94
14. Moberg	5104.96	-47.31	4998.99	-48.41
15. Moberg <sub>one-step</sub>	8779.41	-9.39	7427.09	-23.35
Startwert LAR-Schätzer:				
16. Huber	6738.98	-30.45	7012.69	-27.62
17. Huber <sub>one-step</sub>	7051.91	-27.22	7249.98	-25.17
18. Hampel	5521.98	-43.01	5607.58	-42.12
19. Hampel <sup>1</sup>	6266.95	-35.32	6752.11	-30.31
20. Hampel <sub>one-step</sub>	5861.36	-39.51	6276.77	-35.22
21. McDonald / Newey	9298.11	-4.03	9780.85	0.95
22. Kappenman	9057.94	-6.51	9202.36	-5.02
23. Yuh-AT	5028.72	-48.10	5165.49	-46.69
24. AT <sub>one-step</sub>	5843.01	-39.69	6149.86	-36.53
25. Yuh-ALS	5032.63	-48.06	5275.99	-45.55
26. ALS <sub>one-step</sub>	5693.81	-41.23	6147.92	-36.55
27. Moberg	5080.07	-47.57	4985.52	-48.54
28. Moberg <sub>one-step</sub>	8979.88	-7.32	6805.60	-29.76

<sup>1</sup>: Hampel-Schätzer mit Tuningkonstanten  $c \cdot MAD$  mit  $c = 3.8$ .



Tabelle 2: Schätzfehler in Mill. SFr. und Prozent des Hochrechnungswertes der desaggregierten Bauinvestitionen für das Jahr 1986.

	Mit Schichtenbildung	Ohne Schichtenbildung
Schätzverfahren	Mittlerer Prognosefehler	Mittlerer Prognosefehler
1. KQ	2674.49 (25.20)	3193.81 (30.06)
2. LAR	678.07 (12.62)	1014.56 (17.21)
3. Huber	1294.57 (19.21)	1771.56 (25.26)
4. McDonald / Newey	2289.54 (24.62)	2796.35 (28.59)
5. Kappenman	1915.46 (21.15)	2137.92 (23.23)
6. Yuh-AT	572.48 (11.37)	851.98 (16.49)
7. Yuh-ALS	619.57 (12.22)	845.35 (16.02)
8. Moberg	956.69 (18.74)	1075.78 (21.52)
9. Hampel	1423.75 (22.22)	1438.32 (22.57)
10. Hampel <sup>1</sup>	1833.48 (22.51)	1963.24 (23.46)
11. Huber <sub>one-step</sub> <sup>2</sup>	2956.25 (23.94)	2857.62 (25.88)
12. Huber <sub>one-step</sub> <sup>3</sup>	1794.35 (25.44)	2147.54 (29.62)
13. Hampel <sub>one-step</sub> <sup>2</sup>	2750.88 (21.77)	3196.94 (26.50)
14. Hampel <sub>one-step</sub> <sup>3</sup>	1441.13 (24.59)	1678.64 (26.74)
15. Yuh-AT <sub>one-step</sub> <sup>2</sup>	1886.67 (28.70)	2161.51 (31.32)
16. Yuh-AT <sub>one-step</sub> <sup>3</sup>	1351.78 (23.13)	1706.46 (27.75)
17. Yuh-ALS <sub>one-step</sub> <sup>2</sup>	1774.88 (19.45)	1914.85 (24.69)
18. Yuh-ALS <sub>one-step</sub> <sup>3</sup>	1399.76 (24.58)	1691.64 (32.06)
19. Moberg <sub>one-step</sub> <sup>2</sup>	2631.63 (29.97)	2451.11 (33.00)
20. Moberg <sub>one-step</sub> <sup>3</sup>	2350.18 (26.17)	2110.87 (31.02)

<sup>1</sup>: Hampel-Schätzer mit Tuningkonstanten  $c \cdot MAD$  mit  $c = 3.8$ .

<sup>2</sup>: KQ-Schätzer als Startwert.

<sup>3</sup>: LAR-Schätzer als Startwert.

Aus den beiden Tabellen ist ersichtlich, daß sowohl bei den Hochrechnungen als auch bei den Prognosefehlern zum Teil erhebliche Unterschiede festzustellen sind.

Das Kleinst-Quadrat-Verfahren überschätzt die Investitionen und hat einen relativ großen Prognosefehler. Vor allem bei den Bauinvestitionen ist dieses zu beobachten, denn dort liegt ein diskretionäres Verhalten vor allem bei den Kleinbetrieben vor. Entweder investieren diese, gemessen an den Beschäftigtenzahlen, sehr viel oder geben an, keine Investitionen getätigt zu haben. Diese Ausreißer der Stichprobe als repräsentativ für die Grundgesamtheit zu betrachten, wäre aber falsch. Gerade gegen diese Ausreißer bietet das Kleinst-Quadrat-Schätzverfahren aber keinen Schutz, so daß dadurch die Überschätzung und die hohe Variabilität zu begründen sind.

Bei dem LAR- und den voll iterierten robusten Schätzverfahren werden die Bauinvestitionen unterschätzt. Die Prognosefehler betragen etwas über 20%, was auch nahezu den Ergebnissen von Klein (1989) entspricht, der die Schätzfehler mittels asymptotischen Überlegungen ermittelte.

Bei den Ausrüstungsinvestitionen liefern die voll iterierten robusten Verfahren

die besten Hochrechnungsergebnisse. Der Huber-Schätzer liegt hier nur gering unter dem Sollwert, und der Hampel-Schätzer mit der Tuningkonstanten  $c \cdot MAD$  mit  $c = 3.8$  trifft diesen Wert sogar nahezu exakt.

Sämtliche adaptierende Schätzverfahren, bis auf die nach Kappenman und nach McDonald und Newey, unterschätzen die Investitionen teilweise erheblich. Offensichtlich sind die vorgeschlagenen Gewichtsfunktionen zu restriktiv. Hierauf deuten auch die relativ geringen Prognosefehler hin.

Auch bei den Ausrüstungsinvestitionen ist diese Tendenz zu beobachten: Starke Unterschätzungen der Investitionen gehen einher mit kleinen Prognosefehlern. Bei den aggregierten Daten werden dagegen mit diesen Verfahren hervorragende Ergebnisse erzielt. Vor allem, wenn keine Schichtenbildung berücksichtigt wird, werden die Sollwerte mit diesen Verfahren fast genau prognostiziert.

Die Schätzverfahren nach Kappenman und McDonald/Newey weisen sehr gute Hochrechnungsergebnisse auf. Die Abweichungen mit maximal 6.51% vom Sollwert sind gering. Für diese guten Hochrechnungsergebnisse müssen allerdings beachtliche Prognosefehler mit bis zu nahezu 29% des Hochrechnungsergebnisses in Kauf genommen werden. Insgesamt ist diese Tendenz bei diesen beiden Verfahren stets zu beobachten gewesen. Vor allem bei den desaggregierten Ausrüstungsinvestitionen werden mittels dieser beiden Verfahren exzellente Prognosen getätigt.

Bei den Ein-Schritt-Schätzverfahren werden sowohl hervorragende als auch sehr schlechte Hochrechnungsergebnisse erzielt. Eine eindeutige Tendenz ist nicht erkennbar. In den meisten Fällen aber werden mit den adaptierenden Ein-Schritt-Verfahren die Investitionen unterschätzt. Bei den robusten Ein-Schritt-Schätzern treten sowohl Unter- als auch Überschätzungen auf. Es sind hier zum Teil extrem große Schätzfehler festzustellen.

## 7. Wahl des Startwertes

Eine weitere zu beantwortende Frage stellt sich bezüglich der Diskussion über die Wahl eines geeigneten Startwertes. Ist der LAR-Schätzer dem KQ-Schätzer vorzuziehen oder nicht?

Tabelle 3: Durchschnittliche Iterationszahlen zur Bestimmung des Schätzwertes bei den voll iterierten Verfahren.

Verfahren	Anzahl Iterationen
Startwert: KQ-Schätzer	
Huber	15
Hampel	16
Yuh	112
Moberg	226
Startwert: LAR-Schätzer	
Huber	16
Hampel	17
Yuh	73
Moberg	143

Die Analysen haben ergeben, daß bei den voll iterierten adaptierenden Schätzverfahren sich diese Wahl lediglich in der Anzahl der Iterationen niederschlägt. Auf das Hochrechnungsergebnis hat sie kaum Auswirkungen. Die Anzahl der Iterationen ist mit dem LAR- deutlich geringer als mit dem KQ-Startschätzer. Bei den voll iterierten robusten Verfahren zeigt die Wahl des Startwertes beim Verfahren nach Huber kaum Wirkung. Das Hochrechnungsergebnis variiert nicht in Abhängigkeit der Wahl des Startschätzers und die Iterationszahlen sind auch nahezu identisch.

Beim Verfahren nach Hampel zeigt sich die Wirkung kaum in der Anzahl der Iterationen, dafür umso mehr im Hochrechnungsergebnis. Algorithmen mit "re-descending" Gewichtsfunktionen, zu denen auch das Verfahren nach Hampel zählt, gelangen allgemein zu unterschiedlichen Konvergenzpunkten. Bei der Frage, welcher Schätzwert nun verwendet werden soll, wird in der Literatur vorgeschlagen, den Wert zu verwenden, der mittels des LAR-Startschätzers ermittelt wurde.<sup>4</sup> Bei der Analyse der Investitionsdaten ist aber der KQ-Startschätzer vorzuziehen, denn damit sind bessere Hochrechnungsergebnisse erzielt worden, wogegen die Investitionen mit dem LAR-Startschätzer unterschätzt worden sind.

Bei den Ein-Schritt-Schätzern hat die Wahl des Startschätzers große Auswirkungen. Trotz der Verwendung des 0-Schritt-Schätzers nach Yuh und Hogg bei den adaptierenden Verfahren haben sich große Unterschiede in den Hochrechnungsergebnissen ergeben. Die Hochrechnungen mit dem KQ-Startschätzer sind höher ausgefallen, so daß die Unterschätzung der Investitionen dadurch ein wenig geringer wurde.

Insgesamt läßt sich feststellen, daß der LAR-Schätzer bei den voll iterierten adaptierenden Verfahren dem KQ-Schätzer vorzuziehen ist. Ansonsten hat sich

---

<sup>4</sup>Vgl. zum Beispiel Klein (1985, S. 35).

aber der KQ-Schätzer als Startwert als durchaus konkurrenzfähig erwiesen, trotz seiner bekannten negativen Eigenschaften bezüglich Ausreißern.

## 8. Zusammenfassung

Die vorliegenden Untersuchungen haben insgesamt keine voll befriedigende Hochrechnungsmethode der Investitionen aufzeigen können.

Die adaptierenden Verfahren als Weiterentwicklung (und damit Verbesserung) der robusten Verfahren konnten den Vorteil der Klassifizierung der Residuenverteilung nicht im Hinblick auf ein besseres Hochrechnungsergebnis nutzen. Sie lieferten sogar schlechtere Ergebnisse.

Zwar sind grundsätzlich die adaptierenden Verfahren durch diese Klassifizierungsmöglichkeit den robusten Verfahren vorzuziehen (robuste Verfahren haben zum Beispiel keine Möglichkeit, sich auf schiefe Residuenverteilungen einzustellen), das Problem dieser Verfahren liegt jedoch in der Wahl einer geeigneten Gewichtsfunktion.

Eine Gewichtsfunktion für eine bestimmte Klasse von Verteilungen kann nicht immer Gültigkeit für alle Datensätze besitzen. Wenn das Klassifikationsschema zum Beispiel die Wahl "schwere Tails" trifft, so wird durch die dann angewandte Gewichtsfunktion unterstellt, daß die Verteilung immer gleich schwere Tails hat, bzw. aufgrund der Unsicherheit, wie stark nun die Tails wirklich sind, wird eine sehr robuste Gewichtsfunktion gewählt, um somit einer Überschätzung aus dem Weg zu gehen.

Liegt nun aber eine Verteilung vor, die zwar als eine mit "schweren Tails" zu klassifizieren ist, in dieser Klasse aber relativ leichte Tails aufweist, so tritt eine deutliche Unterschätzung auf. Dies zeigen die vorliegenden Untersuchungen deutlich.

## Literatur

- ANSCOMBE, F. J. (1967), Topics in the Investigation of Linear Relations Fitted by Methods of Least Squares, *Journal of the Royal Statistical Society, Series B*, S. 1–52.
- BÖNTE, G. (1994), *Evaluierung adaptiver Hochrechnungsverfahren*, Diplomarbeit an der Christian-Albrechts-Universität zu Kiel.
- BUTLER, R. J., McDONALD, J. B. UND NELSON, R. D. (1990), Robust and Partially Adaptive Estimation of Regression Models. *The Review of Economics and Statistics*, 72, S. 321–327.
- EFRON, B. UND TIBSHIRANI, R. (1986), Bootstrap Methods For Standard Errors, Confidence Intervals, and Other Methods of Statistical Accuracy, *Statistical Science*, 1, S. 54–77.

- HOGG, R. V. UND RANGLES, R. H. (1972), On the Selection of the Underlying Distribution and Adaptive Estimation. *Journal of the American Statistical Association*, 67, S. 597–600.
- HOGG, R. V. UND RANGLES, R. H. (1975), Adaptive Distribution-Free Regression Methods and their Applications. *Technometrics*, 17, S. 399–407.
- HOGG, R. V. UND RUSSEL, V. L. (1983), On Adaptive Regression, *Journal of Statistical Planning and Inference*, 9, S. 333–343.
- HSIEH, D. A. UND MANSKI, CH. F. (1987), Monte Carlo Evidence on Adaptive Maximum Likelihood Estimation of a Regression. *The Annals of Statistics*, 15, S. 541–551.
- HUBER, P. J. (1973), Robust Regression: A Review, *Annals of Mathematical Statistics*, 43, S. 1041–1067.
- HUBER, P. J. (1977), Robust Methods of Estimation of Regression Coefficients, *Mathematische Operationsforschung und Statistik, Series Statistics*, 8, S. 41–53.
- KAPPENMAN, R. F. (1988), Robust Symmetric Distribution Location Estimation and Regression, *Journal of Statistical Planning and Inference*, 19, S. 55–72.
- KLEIN, I. (1985), *Robuste Statistische Verfahren, Lokalisationsschätzung*, Unveröffentlichtes Manuskript zur Vorlesung "Robuste Verfahren, Lokalisationsschätzung" an der Christian-Albrechts-Universität zu Kiel.
- KLEIN, I. (1989), Hochrechnung der KOF/ETH-Investitionsdaten, *Monatsbericht der KOF/ETH*, Nr. 5.
- LI, GUOYING (1985), Robust Regression, in: Hoaglin, D., Mosteller, F. und Tukey, J. W.: *Exploring Data, Tables, Trends and Shapes*, Wiley & Sons, S. 281–314.
- MCDONALD, J. B. UND WHITE, S. B. (1993), A Comparison of Some Robust, Adaptive and Partially Adaptive Estimators of Regression Models, *Econometric Reviews*, 12, S. 103–124.
- MCDONALD, J. B. UND NEWAY, W. K. (1988), Partially Adaptive Estimation of Regression Models via the Generalized T Distribution, *Econometric Theory*, 4, S. 428–457.
- MOOD, A., GRAYBILL, F. A. UND BOES, D. C. (1974), *Introduction To the Theory of Statistics*, McGraw-Hill International Editions.

- MOBERG, T. F., RANDLES, R. H. UND RAMBERG J. S. (1980), An Adaptive Multiple Regression Procedure Based on M-Estimators, *Technometrics*, 22, S. 213–224.
- MOBERG, T. F., RANDLES, R. H. UND RAMBERG J. S. (1978), An Adaptive M-Estimator and its Application to a Selection Problem. *Technometrics*, 20, S. 255–263.
- TIBSHIRANI, R. (1988), Variance Stabilization and the Bootstrap, *Biometrika*, 75, S. 433–444.
- TUKEY, J.K. (1960), A Survey of Sampling from Contaminated Distributions, in: Olkin, J. (Hrsg.), *Contributions To Propability and Statistics*, Stanford, S. 448–485.
- WERNECKE, K.–D. (1977), Jackknife, Bootstrap und Cross-Validation: Eine Einführung in Methoden der wiederholten Stichprobenziehung, *Allgemeines Statistisches Archiv*, 77, S. 32–59.
- YUH, L. UND HOGG, R. V. (1988), On Adaptive M-Regression, *Biometrics*, 44, S. 433–445.